Miscellanea

Biometrika (1979), **66**, 1, *pp*. 153–5
Printed in Great Britain

Some power studies of a portmanteau test of time series model specification

By NEVILLE DAVIES
Department of Mathematical Sciences, Trent Polytechnic, Nottingham

AND PAUL NEWBOLD
Department of Mathematics, University of Nottingham

Summary

In this note we present simulation evidence on the power of a portmanteau statistic used to detect time series model misspecification. This is related to the loss in forecasting accuracy resulting from use of the incorrectly specified model. Our conclusion is that the statistic achieves a high level of success only when sample size is large.

Some key words: Diagnostic checking; Portmanteau test; Specification error; Time series analysis.

1. Introduction

Suppose that a time series X_t is generated by the stationary ARMA (p, q) process

$$(1-\phi_1 B - \ldots - \phi_p B^p) X_t = (1-\theta_1 B - \ldots - \theta_q B^q) a_t, \tag{1·1}$$

where $B^j X_t = X_{t-j}$ and a_t is zero-mean white noise. An integral part of the methodology of Box & Jenkins (1970) for fitting models (1·1) involves 'diagnostic checks' on the adequacy of representation of an initially identified model to a series of n observations. One such check, developed by Box & Pierce (1970), contemplates general alternatives within the autoregressive-moving average class of models. Denote the residuals from the fitted model by $\hat{a}_t$, with autocorrelations

$$\hat{r}_k = \sum_{t=k+1}^{n} \hat{a}_t \hat{a}_{t-k} \Big/ \sum_{t=1}^{n} \hat{a}_t^2 \quad (k = 1, 2, \ldots).$$

Box & Pierce show that, under the hypothesis of correct model specification, provided that m is moderately large, the statistic

$$Q = n \sum_{k=1}^{m} \hat{r}_k^2 \tag{1·2}$$

is asymptotically distributed as χ^2 with $(m-p-q)$ degrees of freedom. Tests of model adequacy based on this statistic are generally called portmanteau tests.

It has been shown by Davies, Triggs & Newbold (1977) that, for sample sizes commonly found in practice, the actual significance levels of Q can be considerably lower than those predicted by asymptotic theory. However, a simple modification, studied in detail by Ljung & Box (1978),

$$Q' = n(n+2) \sum_{k=1}^{m} (n-k)^{-1} \hat{r}_k^2, \tag{1·3}$$

appears to have a distribution very much closer to the asymptotic χ^2. It would seem preferable, then, to base tests of model adequacy on (1·3) rather than (1·2), and in the remainder of this paper we shall concentrate on the behaviour of the modified statistic Q', and in particular investigate the frequency with which it detects misspecification, relating this to the increase in forecast error variance resulting from use of the incorrect model.

2. Some empirical powers

Let X_t be generated by the process (1·1).

We consider the case where the assumed model is pure autoregressive, of order p^*,

$$\phi^*(B)X_t = \eta_t. \tag{2·1}$$

For fitting such models to data, it will not matter asymptotically whether we use maximum likelihood, least squares or the Yule–Walker equations. It is convenient to consider the latter, in which case the estimates of the coefficients in (2·1) are obtained by solving

$$r_k = \sum_{j=1}^{p^*} \hat{\phi}_j^* r_{k-j} \quad (k = 1, \ldots, p^*), \tag{2·2}$$

where here the r_j are the sample autocorrelations of the data. Now, since sample autocorrelations are consistent estimates of the corresponding population quantities, the probability limits of the $\hat{\phi}_j^*$ are obtained by substituting ρ_j for r_j in (2·2), where the ρ_j are the autocorrelations of the true process, and can most conveniently be derived from (1·1) using an algorithm of McLeod (1975, 1977).

Table 1. *Empirical powers of the statistic Q', with sample size n, for 5% and 10% levels of significance*

Model				AR	% loss	n = 50		n = 100		n = 200	
ϕ_1	ϕ_2	θ_1	θ_2			5%	10%	5%	10%	5%	10%
0	0	−0·2	0·4	4	1·3	0·059	0·102	0·083	0·142	0·119	0·205
0·3	0	0·75	0	4	1·5	0·055	0·110	0·077	0·152	0·104	0·173
0·9	0	−0·25	0	1	5·5	0·218	0·279	0·270	0·361	0·482	0·597
0	0	−0·6	−0·4	1	5·6	0·170	0·238	0·267	0·360	0·457	0·563
0·8	0	0·2	0·4	1	5·9	0·161	0·243	0·239	0·343	0·438	0·559
0·8	0	0·6	0·4	4	7·3	0·057	0·104	0·087	0·155	0·172	0·261
0·8	−0·4	−0·8	0	4	9·3	0·174	0·282	0·241	0·367	0·509	0·668
1·6	−0·9	−0·8	0	4	12·3	0·728	0·810	0·837	0·891	0·916	0·956
0·4	0	−0·2	−0·4	1	15·8	0·250	0·331	0·348	0·457	0·688	0·795
0·3	0	1	0	4	17·1	0·107	0·174	0·195	0·315	0·445	0·612
0	0	−0·6	0·4	4	17·9	0·103	0·177	0·237	0·359	0·494	0·663
0	0	−0·2	0·4	1	18·8	0·330	0·443	0·592	0·700	0·934	0·972
0·6	0	1	0	1	20·0	0·166	0·244	0·266	0·397	0·506	0·659
0·4	0	1·4	−0·4	4	20·0	0·139	0·222	0·229	0·365	0·566	0·721
0	0	−0·9	−0·8	1	37·9	0·439	0·548	0·789	0·874	0·993	1·000
0	0	−1·2	−1	4	38·5	0·239	0·365	0·539	0·678	0·955	0·986
0·8	0	0·2	0·8	1	39·9	0·416	0·533	0·759	0·852	0·994	0·998
0·6	0	−0·75	0	1	41·6	0·508	0·616	0·859	0·922	0·999	1·000
0·4	0	1·8	−0·8	4	41·9	0·203	0·304	0·454	0·580	0·903	0·958
0·6	0	−1	0	1	80·0	0·582	0·703	0·911	0·957	1·000	1·000
0	0	−0·4	−1	1	86·4	0·588	0·700	0·955	0·982	1·000	1·000
0	0	−2	−1	4	86·7	0·280	0·428	0·611	0·748	0·983	0·995
0·4	0	1·8	−0·8	1	106	0·568	0·695	0·902	0·958	1·000	1·000
0·8	0	−2	−1	4	107	0·583	0·700	0·810	0·906	0·999	1·000

AR denotes the order of the autoregressive model fitted.

% loss denotes the percentage increase in expected squared forecast error one step ahead resulting from use of the misspecified model.

Data were generated from models of the form $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$.

An estimate of the percentage increase in forecast error variance, subsequently termed percentage loss, from use of the misspecified rather than the correct model is then found using results given by N. Davies in a Nottingham Ph.D. thesis and by the present authors in an unpublished report.

We simulated data from a number of autoregressive-moving average models and estimated autoregressive models of order either one or four. Note that use of the autoregressive misspecification was dictated by the relatively low costs of estimating autoregressive rather than moving average or mixed models. The examples were chosen so as to provide a wide range of percentage losses for forecasting one step ahead with the misspecified model. All simulation experiments were based on 1000 replications, and the proportion of times that misspecification was detected for tests at the 5% and 10% levels of significance was recorded. These powers are recorded for the modified statistic Q' of (1·3) in Table 1. In all cases m was fixed at 20.

Table 1 presents rather a mixed picture. For a sample size of 50, the portmanteau test fails to detect misspecification disturbingly often, even when the consequences of proceeding with the misspecified model in terms of forecast accuracy are very severe. However, as is to be expected as sample size increases, the performance of the test improves dramatically, so that for 200 observations, the probability of correctly rejecting the null hypothesis of adequate specifications is typically very high in those cases for which misspecification leads to a large increase in one-step-ahead forecast error variance.

In assessing Q' from Table 1 it might be argued that the times the test does detect a model misspecification are just the times the model would have predicted badly. To examine this possibility, models were chosen from Table 1, series of length 51 and 101 were generated from them with AR(1) models being estimated in each case from the first 50 and 100 observations respectively. For all models chosen the mean squared prediction error of 51st and 101st observations, using the fitted autoregressions, did not differ appreciably between cases when Q' was or was not significant. Hence, we can sensibly assess the portmanteau test via the one-step-ahead loss of forecasting precision.

References

Box, G. E. P. & Jenkins, G. M. (1970). *Time Series Analysis Forecasting and Control.* San Francisco: Holden Day.

Box, G. E. P. & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Am. Statist. Assoc.* **65**, 1509–26.

Davies, N., Triggs, C. M. & Newbold, P. (1977). Significance levels of the Box–Pierce portmanteau statistic in finite samples. *Biometrika* **64**, 517–22.

Ljung, G. M. & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303.

McLeod, I. (1975). Derivation of the theoretical autocovariance function of autoregressive-moving average time series. *Appl. Statist.* **24**, 255–6. Correction (1977), *Appl. Statist.* **26**, 194.

[*Received April* 1978. *Revised September* 1978]